

---

**Complete nucleotide sequence of the influenza A/PR/8/34 virus NS gene and comparison with the NS genes of the A/Udorn/72 and A/FPV/Rostock/34 strains**

---

**M.Baez, R.Taussig, J.J.Zazra, J.F.Young and P.Palese**

Department of Microbiology, Mount Sinai School of Medicine of CUNY, One Gustave Levy Place, New York, NY 10029, and

**A.Reisfeld and A.M.Skalka**Roche Institute of Molecular Biology, Nutley, NJ 07110, USA

---

Received 16 September 1980

---

**ABSTRACT**

The nucleotide sequence of the NS gene of the human influenza virus A/PR/8/34 was determined and found to be the same length (890 nucleotides) as the NS gene of another human influenza virus A/Udorn/72 and of the avian isolate A/FPV/Rostock/34. Comparison of the sequences of the NS genes of the two human influenza viruses shows an 8.9% difference whereas the NS gene of the avian isolate differs by only 8% from that of the human strain A/PR/8/34. The extensive sequence similarity among these three genes does not support the notion of species specific homology groups among NS genes of avian and human influenza virus strains. The primary sequence of the A/PR/8/34 NS gene is consistent with the finding that the influenza virus NS gene may code for two overlapping polypeptides. In addition, an open reading frame potentially coding for a polypeptide 167 amino acids in length was found in the negative strand RNA of the A/PR/8/34 virus NS gene.

**INTRODUCTION**

Influenza A viruses contain eight single stranded RNA segments (1-4) each of which encodes at least one polypeptide (5-10). The genome of these viruses has been shown to undergo continuous variation for which several mechanisms have been identified. These include reassortment among the genes of human influenza viruses, reassortment involving animal influenza virus strains (11, 12, 13) and sequential mutation in individual genes of viruses of the same subtype (14). Until recently, direct comparison of influenza virus gene sequences was limited to RNA-RNA hybridization and oligonucleotide mapping techniques. With the availability of rapid DNA cloning and sequencing techniques, the absolute differences among the genes of different viruses can be examined and the extent of variation determined. Furthermore,

one can use comparative nucleotide sequence analysis to determine the structural basis for gene expression, regulation and packaging. Within the last year, the nucleotide sequences of several hemagglutinin genes (15-20) and one matrix protein gene (21) have been reported and most recently the nucleotide sequences of the nonstructural (NS) protein genes obtained from two different influenza viruses, one avian isolate, A/FPV/Rostock/34 (Hav1N1) (FPV), and one human isolate, A/Udorn/72 (H3N2) (Udorn), have been reported (22,23). Results from the analyses of these NS gene sequences are consistent with previous genetic studies (9,10, 24,25) which showed that the NS gene codes for two overlapping polypeptides. Lamb and Lai (23) also showed that the NS2 polypeptide is most likely translated from a spliced mRNA whose intervening sequence is contained in the coding region of the NS1 gene.

We report here the complete nucleotide sequence of the NS gene obtained from a third strain, A/PR/8/34 (H0N1) (PR8), a human isolate. We find that the NS gene of the PR8 virus has a nucleotide sequence which is consistent with the coding capacity and structural organization reported for the FPV (22) and Udorn (23) virus NS genes. In comparing the PR8 NS gene to that of the FPV and Udorn viruses we observe greater than 84% homology among the nucleotide and amino acid sequences of all three genes. This last observation is in contrast to previous findings based on RNA-RNA hybridization which places the NS gene of the PR8 virus and those of other human influenza viruses in a different homology group from that of the NS gene of the FPV virus (26).

### MATERIALS AND METHODS

#### Virus

The human influenza virus A/PR/8/34 (H0N1) was grown in embryonated hen's eggs. Virus purification and RNA extraction have been described (27).

#### Bacterial strains

E. coli strains DP50/supF and C600 were used to propagate recombinant  $\lambda$  phage and plasmid pBR322 respectively.

#### Cloning into $\lambda$ gtWES- $\lambda$ B

Double stranded cDNA transcripts of the PR8 viral RNAs were

---

prepared as described (28) except that a synthetic oligonucleotide, d(AGCAAAAGCAG)rG (Collaborative Research, Waltham, MA), which is complementary to the 3' end of all eight RNAs (29, 30, 31) was used to prime first strand synthesis. After digestion of the double stranded DNAs with S1 nuclease, the blunt ended DNAs were methylated with EcoRI methylase and EcoRI linkers (Collaborative Research) were added (32). Following digestion with EcoRI endonuclease, the cDNAs were ligated to the left and right arms of  $\lambda$ gtWES $\cdot\lambda$ B and recombinant phages were produced by in vitro packaging of the chimeric DNAs (33). E. coli DP50/supF cells were infected with phage and the resulting plaques were examined for influenza virus gene sequences by in situ hybridization (34) using as probe viral RNA which was partially digested with alkali and 5' end labeled with  $\gamma$ -<sup>32</sup>P-ATP and polynucleotide kinase. The gene derivation of the viral sequences in each recombinant phage was identified by the "Northern" blot technique (35): Viral RNA was fractionated by agarose gel electrophoresis and the RNA blotted onto diazobenzylxymethyl paper (Schleicher and Schuell, Inc.). Recombinant phage DNA was nick translated (36) and hybridized to the immobilized RNA. The gene derivation was determined by the electrophoretic mobility of the hybridizing RNA segment (5). One NS gene specific insert from a  $\lambda$ gtWES clone was subcloned into pBR322 (37) and designated pAR109.

#### Direct cloning into pBR322

Double stranded cDNA transcripts of PR8 virus RNA were made as described above except that the second strand synthesis was primed with a second synthetic oligonucleotide, d(AGTAGAAACAAG) (Collaborative Research), which is complementary to the first twelve residues of the 3' end of the positive cDNA strand (29, 30, 31). HindIII linkers (Collaborative Research) were added to the DNAs with T4 DNA ligase (generously provided by Dr. Kathy Squires, Columbia University, New York) and the resulting DNA was digested with HindIII endonuclease (32). DNA fragments, fractionated on a 1.5% low melting point agarose gel (Bethesda Research Laboratories), were visualized by ethidium bromide staining, excised, extracted and ligated into the HindIII site of pBR322 (37). E. coli strain C600 was transformed with the

chimeric DNA by the calcium chloride technique (38), and colony hybridization with an  $\alpha$ -<sup>32</sup>P-dCTP labeled viral cDNA probe was used to identify bacterial clones carrying influenza virus specific DNA (39). The gene sequences contained in each recombinant plasmid were identified by the "Northern" blot technique described above. One of the NS gene specific clones (pJZ101) was used for sequence analysis.

### DNA sequencing

Three sequencing techniques were used to determine the sequence of the PR8 NS gene: the chemical method of Maxam and Gilbert (40); the Forward-Backward method of Seif et al. (41); and the chain termination procedure of Sanger et al. (42) using restriction fragments of cloned DNA to prime DNA synthesis from viral RNA with reverse transcriptase (16) (generously provided by Dr. J.W. Beard, Division of Cancer Cause and Prevention, National Cancer Institute).

Nucleotide sequence data were stored, edited and analyzed in a DEC PDP11/45 digital computer using published programs (43, 44, 45).

### RESULTS

Initial attempts to clone the NS gene of the PR8 virus resulted in clones which contained inserts less than 550 base pairs long and could not be representative of the entire NS gene. One of these clones, pAR109, was used for further analysis. Subsequently, a second approach using the synthetic primer d(AGTAGAAACAAG) which is complementary to the 3' terminal sequences of full length positive strand cDNA (29, 30, 31) resulted in NS gene specific clones with DNA inserts of more than 800 base pairs. The NS gene specific DNAs from one of the latter clones (pJZ101) and from pAR109 were screened with several restriction endonucleases and restriction maps were constructed. DNA sequencing of the ends of these clones and comparison to direct RNA sequencing of the PR8 virus NS gene (29, 30, 31) showed that the pAR109 insert is 530 base pairs long and lacks the first 23 residues found at the 3' terminus of the NS gene vRNA. The pJZ101 insert was found to be identical to the 5' end of the NS vRNA and

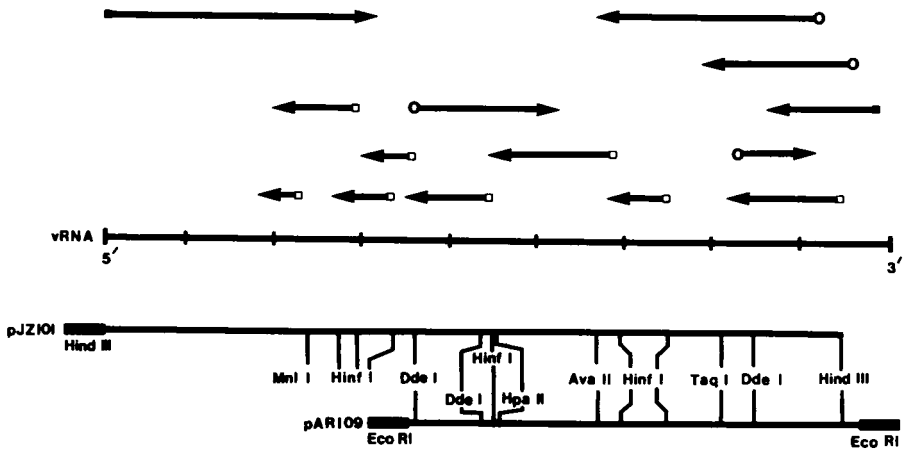


Figure 1. Sequencing strategy and restriction endonuclease maps of the NS gene inserts. The DNA inserts from clones pJZ101 and pAR109 are shown with their restriction endonuclease sites and are aligned with respect to the PR8 NS gene. The large closed rectangles represent the synthetic HindIII and EcoRI linkers at the ends of the cloned DNAs. Arrows indicate the direction of sequencing and the length of the sequence determined by a particular technique; (O) the Forward-Backward technique of Seif et al. (41), (■) the chemical method of Maxam and Gilbert (40), (□) the chain termination technique of Sanger et al. (16, 42).

it extends 843 residues to the internal HindIII site at a position 47 nucleotides from the 3' end of the vRNA.

The strategy used to determine the nucleic acid sequence of the PR8 virus NS gene is shown in figure 1 and the complete nucleotide sequence and predicted amino acid sequences of the viral polypeptides are presented in figure 2. The first initiation codon in the PR8 virus NS gene is found at positions 27 to 29 and

Figure 2. (Overleaf) Complete nucleotide and predicted amino acid sequences of the PR8 virus NS gene and comparison to those of the NS genes of the A/FPV/Rostock/34 and A/Udorn/72 viruses. Nucleotide differences among the sequences are underlined, and amino acid changes are indicated by rectangles. The arrowheads indicate the possible splicing sites for the NS2 genes. The NS gene sequences of the A/FPV/Rostock/34 and A/Udorn/72 viruses are taken from references 22 and 23 respectively.

		NB1 & NB2	(NB1 →)
		30	60
A/FPV/NDSTOCK/34	AGCAAAAGCAGGGTGCACAAACATA	ATG GAT TCC AAC ACT GTG TCA AGC TTT CAG GTA GAC	↑
		MET ASP SER ASN THR VAL SER SER PHE GLN VAL ASP	
A/PR/8/34	AGCAAAAGCAGGGTGCACAAACATA	ATG GAT CCA AAC ACT GTG TCA AGC TTT CAG GTA GAI	↑
		MET ASP PRO ASN THR VAL SER SER PHE GLN VAL ASP	↑
A/UDORN/72	AGCAAAAGCAGGGTGCACAAACATA	ATG GAT TCC AAC ACT GTG TCA AGI TTT CAG GTA GAC	↑
		MET ASP SER ASN THR VAL SER SER PHE GLN VAL ASP	↑

TGC TTT CTT TGG CAT GTC CGC AAA CGA GTT GCA GAC CAA GAA ATE GGT GAT GCC CCA TTC CTT GAC CGA CTT CGC  
 CYS PHE LEU TRP HIS VAL ARG LYS ARG PHE ALA ASP GLN GLU MET GLY ASP ALA PRO PHE LEU ASP ARG LEU ARG

TGC TTT CTT TGG CAT GTC CGC AAA CGA GTT GCA GAC CAA GAA CTA GGT GAT GCC CCA TTC CTT GAT CGG CTT CGC  
 CYS PHE LEU TRP HIS VAL ARG LYS ARG VAL ALA ASP GLN GLU LEU GLY ASP ALA PRO PHE LEU ASP ARG LEU ARG

TGC TTC CTT TGG CAT GTC CGA AAA CAA GTT GTA GAC CAA GAA CTA GGT GAT GCC CCA TTC CTT GAT CGG CTT CGC  
 CYS PHE LEU TRP HIS VAL ARG LYS GLN VAL VAL ASP GLN GLU LEU GLY ASP ALA PRO PHE LEU ASP ARG LEU ARG

CGA GAT CAG AAG TCC CTG AAG GGA AGA GGC AGC ACT CTT GGT CTG GAC ATC GAC ACA GCI ACI CGT GTT GGA ARG  
 ARG ASP GLN LYS SER LEU ARG GLY ARG GLY SER THR LEU GLY LEU ASP ILE ASP THR ALA THR ARG VAL GLY LYS

CGA GAT CAG AAG TCC CTA AGA GGA AAG GGC AGC ACC CTC GGT CTG GAC ATC GAG ACA GCC ACA CGT GCT GGA ARG  
 ARG ASP GLN LYS SER LEU ARG GLY ARG GLY SER THR LEU GLY LEU ASP ILE GLU THR ALA THR ARG ALA GLY LYS

CGA GAT CAG AAG TCC CTA AAG GGA AGA GGC AGC ACT CTC GGT CTA AAC ATC GAA GCA GCC ACC CAT GTT GGA ARG  
 ARG ASP GLN LYS SER LEU ARG GLY ARG GLY SER THR LEU GLY LEU ASN ILE GLU ALA ALA THR HIS VAL GLY LYS

CAG ATA GTG GAG CGG ATT CTG CAG GAC GAA TCC GAT GAG GCA CTT AAA ATG ACC ATT GCC TCT GTA CCT GCT ACA  
 GLN ILE VAL GLU ARG ILE LEU GLU ASP GLU SER ASP GLU ALA LEU LYS MET THR ILE ALA SER VAL PRO ALA THR

CAG ATA GTG GAG CGG ATT CTG AAA GAA GAA TCC GAT GAG GCA CTT AAA ATG ACC ATG GCC TCT GTA CCT GCG TCG  
 GLN ILE VAL GLU ARG ILE LEU LYS GLU GLU SER ASP GLU ALA LEU LYS MET THR MET ALA SER VAL PRO ALA SER

CAG ATA GTA GAG AAG ATT CTG AAG GAA GAA TCI GAT GAG GCA CTT AAA ATG ACC ATG GCC TCC ACA CCT GCT TCG  
 GLN ILE VAL GLU LYS ILE LEU LYS GLU GLU SER ASP GLU ALA LEU LYS MET THR MET ALA SER THR PRO ALA SER

CGC TAC CTA ACT GAC ATG ACT CTT GAA GAG ATG TCA AAG GAC TGG TTC ATG CTC ATG CCC AAA CAG AAA GTG GCA  
 ARG TYR LEU THR ASP MET THR LEU GLU GLU MET SER ARG ASP TRP PHE MET LEU MET PRO LYS GLN LYS VAL ALA

CGI TAC CTA ACT GAC ATG ACT CTT GAG GAA ATG TCA AAG GAC TGG TCC ATG CTC ATG CCC AAG CAG AAA GTG GCA  
 ARG TYR LEU THR ASP MET THR LEU GLU GLU MET SER ARG ASP TRP SER MET LEU ILE PRO LYS GLN LYS VAL ALA

CGA TAC ATA ACT GAC ATG ACT ATT GAG GAA ITG TCA AAG GAC TGG TTC ATG CTG ATG CCC AAG CAG AAA GTG GAA  
 ARG TYR ILE THR ASP MET THR ILE GLU GLU LEU SER ARG ASP TRP PHE MET LEU MET PRO LYS GLN LYS VAL GLU

GCC TCC CTT TGC ATC AGA ATG GAC CAG GCG ATC ATG GGG AAG AAC ATC ATA CTG AAA GCA AAC TTC AGT GTG ATT  
 GLY SER LEU CYS ILE ARG MET ASP GLN ALA ILE MET GLY LYS ASN ILE ILE LEU LYS ALA ASN PHE SER VAL ILE

GCC CCT CTT TGT ATC AGA ATG GAC CAG GCG ATC ATG GAT AAG AAC ATC ATA CTG AAA GCG AAC TTC AGT GTG ATT  
 GLY PRO LEU CYS ILE ARG MET ASP GLN ALA ILE MET ASP LYS ASN ILE ILE LEU LYS ALA ASN PHE SER VAL ILE

GGA CCT CTT TGC ATC AGA ATG GAC CAG GCA ATC ATG GAT AAG AAC ATC ATG ITG AAA GCG AAT TTC AGT GTG ATT  
 GLY PRO LEU CYS ILE ARG ILE ASP GLN ALA ILE MET ASP LYS ASN ILE MET LEU LYS ALA ASN PHE SER VAL ILE



the reading frame remains open until the termination codon at nucleotides 717 to 719. This region has a coding capacity for a polypeptide of 230 amino acids. A second reading frame starting at positions 469 to 471 can encode a polypeptide of up to 131 amino acids in length. Lamb and Lai (23) have demonstrated that a second polypeptide, NS2, may be translated from a spliced mRNA, the 5' end of which is the same as the NS1 mRNA. Our sequence contains the same potential splicing signals and is consistent with this gene arrangement. Possible splice signals are found at nucleotide positions 54 to 57 and 526 to 529 of the PR8 NS RNA. This spliced mRNA would code for a polypeptide 121 amino acids long.

The NS gene sequence and predicted amino acid sequences of the PR8 virus are compared to those of the FPV (22) and Udorn (23) viruses in figure 2. The PR8 virus NS gene is 890 nucleotides long and is the same length as that found for the FPV and Udorn virus NS genes. Table 1 lists the percentage of single base differences among the nucleotide sequences of these three genes while table 2 presents the differences deduced for their amino acid sequences.

### DISCUSSION

Examination of the complete sequence reveals that the PR8 virus NS gene is 890 nucleotides long and may code for at least two polypeptides. The structural organization of the NS1 and NS2 genes within the smallest RNA segment of the PR8 virus genome is consistent with the model proposed by Lamb and Lai (23) for that of the Udorn virus NS gene. This model predicts that the mRNA of the PR8 virus NS2 gene may derive from a splicing event which excises sequences that encode a portion of the NS1 gene.

In a study by another group, the sequence of the 3' terminal portion of the PR8 virus NS gene has been determined by chain termination sequencing using a synthetic oligonucleotide to prime cDNA synthesis (46, 47). Among the 126 nucleotides that were reported, there are six differences relative to the PR8 NS gene sequence presented here and these base substitutions would result in two amino acid changes. Clonal variation might



Table 1. DIFFERENCES IN NUCLEOTIDE SEQUENCES AMONG THE NS GENES OF THREE INFLUENZA VIRUSES

	Total (%)	NS1 (%)	NS2 (%)
A/PR/8/34 - A/FPV/Rostock/34	71/890 (8.0)	61/690 (8.8)	24/363 (6.6)
A/PR/8/34 - A/Udorn/72	78/890 (8.9)	63/690 (9.1)	29/363 (8.0)
A/Udorn/72 - A/FPV/Rostock/34	96/890 (10.8)	80/690 (11.6)	31/363 (8.5)

The lengths of the NS genes, NS1 coding regions and NS2 coding regions are 890, 690 and 363 nucleotides, respectively. The differences reported for the NS1 coding regions were calculated only for the "common" 690 nucleotides. It should be noted, however, that the Udorn virus NS1 coding region is 21 nucleotides longer than those of the PR8 and FPV viruses.

Table 2.  
DIFFERENCES IN AMINO ACID SEQUENCES AMONG THE NS POLYPEPTIDES OF THREE INFLUENZA VIRUSES

	NS1 (%)	NS2 (%)
A/PR/8/34 - A/FPV/Rostock/34	24/230 (10.4)	10/121 (8.3)
A/PR/8/34 - A/Udorn/72	26/230 (11.3)	9/121 (7.4)
A/Udorn/72 - A/FPV/Rostock/34	35/230 (15.2)	7/121 (5.8)

The lengths of the NS1 and NS2 polypeptides are 230 and 121 amino acids respectively. The differences reported for the NS1 polypeptides were calculated with respect to the "common" 230 amino acid region. It should be noted, however, that the Udorn virus NS1 polypeptide is seven amino acids longer than those of the PR8 and FPV viruses.

explain the 5% difference between the two PR8 gene sequences; however, in a study that compared the nucleotide sequences of two hemagglutinin genes derived from viruses isolated three years apart, a difference of only 3% was noted (16).

We have compared the complete nucleotide sequences derived from the NS genes of the FPV, PR8 and Udorn viruses and find that all three NS genes are 890 residues long. The small number of differences observed among these NS genes clearly demonstrates that they are closely related and could have derived from a common ancestral gene. A comparison between the NS genes of the PR8 and FPV viruses shows an 8% difference resulting from single base changes. The comparison between the NS genes of the PR8 and Udorn viruses shows a slightly greater difference of 8.9% (see table 1). It appears from this observation that the PR8 virus NS gene is more closely related to the FPV NS gene than to the Udorn NS gene. This is in contrast to previous findings based on RNA-RNA hybridization analysis which showed that the PR8 NS gene is in a homology group distinct from that of the FPV virus (26). The suggestion that RNA-RNA hybridization analysis can be used to group the NS genes of human and avian influenza viruses into species specific homology groups is thus not supported by the available nucleotide sequence data. However, sequencing data from additional human and avian influenza virus strains will be necessary to resolve this point.

We have also compared the amino acid sequences deduced from the three NS gene sequences. The NS1 polypeptides of the human strain PR8 and the avian strain FPV are of equal length (230 amino acids) whereas that of the other human strain Udorn is seven amino acids longer. The NS2 polypeptides of all three strains appear to have the same length of 121 amino acids. The observation that the NS2 polypeptides show fewer amino acid differences (6% to 8%) than do the NS1 polypeptides (10% to 15%) may indicate that the NS2 portions of the NS gene are more faithfully conserved (see table 2).

It is interesting to note that the negative strands of the three NS genes have an initiation codon at a position 98 residues from their 5' ends and an open reading frame that extends to

---

residue 598 in the NS genes of the PR8 and FPV viruses and to residue 745 in the NS gene of the Udorn virus. A polypeptide of 167 amino acids would be predicted for the negative strand of the PR8 and FPV virus NS genes and a polypeptide of 216 amino acids can be predicted for that of the Udorn virus NS gene. The absence of termination codons for such a long distance is unexpected, since the probability of finding an open reading frame of 167 or 216 amino acids is 0.0003 and 0.00003, respectively. (This calculation is based on the random appearance of three termination codons out of the 64 possible triplets.) Conversely, 7-12 termination codons are found in each of the remaining reading frames over the corresponding area of the FPV and PR8 NS genes and similarly 11 and 14 termination codons are observed in the two remaining reading frames of the Udorn NS gene. The significance of the "agnogene" sequences (48) in the influenza virus NS genes is not known but merits further investigation.

### ACKNOWLEDGEMENTS

We thank Dr. Ravi Dhar for helpful suggestions regarding sequencing techniques and we thank Dr. Robert A. Lamb for providing us with sequence data prior to publication. This work was supported in part by Public Health Service Grant AI-11823 from the National Institute of Allergy and Infectious Diseases, by Grant PCM-76-11066 from the National Science Foundation and by Grant MV-23A from the American Cancer Society. P.P. is a recipient of an I.T. Hirschl Career Scientist Award. J.J.Z. is the recipient of an American Lung Association Fellowship.

### REFERENCES

1. Pons, M.W. (1976) *Virology* 69, 789-792.
2. Bean, W.J. and Simpson, R.W. (1976) *J. Virol.* 18, 365-369.
3. Palese, P. and Schulman, J.L. (1976) *J. Virol.* 17, 876-884.
4. McGeoch, D.J., Fellner, P. and Newton, C. (1976) *Proc. Natl. Acad. Sci. USA* 73, 3045-3049.
5. Ritchey, M.B., Palese, P. and Schulman, J.L. (1976) *J. Virol.* 20, 307-313.
6. Scholtissek, C., Harms, E., Rohde, W., Orlich, M. and Rott, R. (1976) *Virology* 74, 332-344.
7. Inglis, S.C., McGeoch, D.J. and Mahy, B.W.J. (1977) *Virology* 78, 522-536.
8. Inglis, S.C., Carroll, A.R., Lamb, R.A. and Mahy, B.W.J. (1976) *Virology* 74, 489-503.
9. Inglis, S.C., Barret, T., Brown, C.M. and Almond, J.W. (1979) *Proc. Natl. Acad. Sci. USA* 76, 3790-3794.

10. Lamb, R.A. and Choppin, P.W. (1979) Proc. Natl. Acad. Sci. USA 76, 4908-4912.
11. Young, J.F. and Palese, P. (1979) Proc. Natl. Acad. Sci. USA 76, 6547-6551.
12. Desselberger, U., Nakajima, K., Alfino, P., Pedersen, F.S., Haseltine, W.A., Hannoun, C. and Palese, P. (1978) Proc. Natl. Acad. Sci. USA 75, 3341-3345.
13. Scholtissek, C., Rohde, W., von Hoyningen, V. and Rott, R. (1978) Virology 87, 13-20.
14. Young, J.F., Desselberger, U. and Palese, P. (1979) Cell 18, 73-83.
15. Porter, A.G., Barber, C., Carey, N.H., Hallewell, R.A., Threlfall, G. and Emtage, J.S. (1979) Nature 282, 471-477.
16. Both, G.W., Sleigh, M.J., Bender, Y.J. and Moss, B.A. (1980) in Proceedings of the International Workshop on Structure and Variation in Influenza Virus in Thredbo, Australia, Laver, W.G. and Air, G.M., Eds., pp. 81-89 Elsevier/North Holland, New York.
17. Brownlee, G.G. (1980) in Proceedings of the International Workshop on Structure and Variation in Influenza Virus in Thredbo, Australia, Laver, W.G. and Air, G.M., Eds., pp. 385-390 Elsevier/North Holland, New York.
18. Gething, M.J., Bye, J., Skehel, J.J. and Waterfield, M. (1980) in Proceedings of the International Workshop of Structure and Variation in Influenza Virus in Thredbo, Australia, Laver, W.G. and Air, G.M., Eds., pp. 1-10 Elsevier/North Holland, New York.
19. Min Jou, W., Verhoeyen, M., Devos, R., Saman, E., Rongxiang, F., Huylebroeck, D., Fiers, W., Threlfall, G., Barber, C., Carey, N. and Emtage, S. (1980) Cell 19, 683-696.
20. Verhoeyen, M., Fang, R., Min Jou, W., Devos, R., Huylebroeck, D., Saman, E. and Fiers, W. (1980) Nature 286, 771-776.
21. Winter, G. and Fields, S. (1980) Nucleic Acids Res. 8, 1966-1974.
22. Porter, A.G., Smith, J.C. and Emtage, J.S. (1980) Proc. Natl. Acad. Sci. USA (in press).
23. Lamb, R.A. and Lai, C.-J. (1980) Cell 21, 475-485.
24. Lamb, R.A., Choppin, P.W., Chanock, R.M. and Lai, C.-J. (1980) Proc. Natl. Acad. Sci. USA 77, 1857-1861.
25. Lamb, R.A., Etkind, P.R. and Choppin, P.W. (1978) Virology 91, 60-78.
26. Scholtissek, C. and von Hoyningen-Huene, V. (1980) Virology 102, 13-20.
27. Palese, P. and Schulman, J.L. (1976) Proc. Natl. Acad. Sci. USA 73, 2142-2146.
28. Emtage, J.S., Catlin, G.H. and Carey, N.H. (1979) Nucleic Acids Res. 6, 1221-1239.
29. Desselberger, U., Racaniello, V.R., Zazra, J.J. and Palese, P. (1980) Gene 8, 315-328.
30. Robertson, J.S. (1979) Nucleic Acids Res. 6, 3745-3757.
31. Skehel, J.J. and Hay, A.J. (1978) Nucleic Acids Res. 5, 1207-1219.
32. Maniatis, T., Hardison, R.C., Lacy, E., Laver, J., O'Connell, C., Quon, D., Sim, G.K. and Efstratiadis, A. (1978) Cell 15, 687-701.

33. Sternberg, N., Tiemeier, D. and Enquist, L. (1977) *Gene* 1, 255-280.
34. Benton, W.D. and Davis, R.W. (1977) *Science* 196, 180-182.
35. Alwine, J.C., Kemp, D.J., Parker, B.A., Reisner, J., Renart, J., Start, G.R. and Wahl, G.M. (1980) in *Methods in Enzymology*, Wu, R., Ed., Vol. 68, pp. 220-242 Academic Press, New York.
36. Maniatis, T., Jeffrey, A. and Kleid, D.G. (1975) *Proc. Natl. Acad. Sci. USA* 72, 1184-1188.
37. Bolivar, F. and Backman, K. (1979) in *Methods in Enzymology*, Wu, R., Ed., Vol. 68, pp. 245-267, Academic Press, New York.
38. Dagert, M. and Ehrlich, S.D. (1979) *Gene* 6, 23-28.
39. Grunstein, M. and Hogness, D.S. (1975) *Proc. Natl. Acad. Sci. USA* 72, 3961-3965.
40. Maxam, A.M. and Gilbert, W. (1980) in *Methods in Enzymology*, Grossman, L. and Moldave, K., Eds., Vol. 65, pp. 499-559, Academic Press, New York.
41. Seif, I., Khoury, G. and Dhar, R. (1980) *Nucleic Acids Res.* 8, 2225-2240.
42. Sanger, F., Nicklen, S. and Coulson, A.R. (1977) *Proc. Natl. Acad. Sci. USA* 74, 5463-5467.
43. Staden, R. (1977) *Nucleic Acids Res.* 4, 4037-4051.
44. Staden, R. (1978) *Nucleic Acids Res.* 5, 1013-1015.
45. Staden, R. (1979) *Nucleic Acids Res.* 6, 2601-2610.
46. Air, G.M. (1980) in *Proceedings of the International Workshop on Structure and Variation in Influenza Virus in Thredbo, Australia*, Laver, W.G. and Air, G.M., Eds., pp. 135-146, Elsevier/North Holland, New York.
47. Air, G.M. and Hackett, J.A. (1980) *Virology* 103, 291-298.
48. Seif, I., Khoury, G. and Dhar, R. (1979) *Cell* 19, 963-977.